

# Memory for musical tempo: Additional evidence that auditory memory is absolute

DANIEL J. LEVITIN

University of Oregon, Eugene, Oregon  
and Stanford University, Stanford, California

and

PERRY R. COOK

Stanford University, Stanford, California

We report evidence that long-term memory retains absolute (accurate) features of perceptual events. Specifically, we show that memory for music seems to preserve the absolute tempo of the musical performance. In Experiment 1, 46 subjects sang two different popular songs from memory, and their tempos were compared with recorded versions of the songs. Seventy-two percent of the productions on two consecutive trials came within 8% of the actual tempo, demonstrating accuracy near the perceptual threshold (JND) for tempo. In Experiment 2, a control experiment, we found that folk songs lacking a tempo standard generally have a large variability in tempo; this counters arguments that memory for the tempo of remembered songs is driven by articulatory constraints. The relevance of the present findings to theories of perceptual memory and memory for music is discussed.

A fundamental problem facing memory theorists is how to account for two seemingly disparate properties of memory. On the one hand, a rich body of literature suggests that the role of memory is to preserve the gist of experiences; memory functions to formulate general rules and create abstract concepts on the basis of specific exemplars (Posner & Keele, 1970; Rosch, 1975). On the other hand, there is the extensive literature suggesting that memory accurately preserves absolute features of experiences (Brooks, 1978; Jacoby, 1983; Medin & Schaffer, 1978). (For a further discussion of these two perspectives, see McClelland & Rumelhart, 1986.)

These perspectives on the function of human memory parallel an old debate in the animal-learning literature about whether animals' internal representations are relational or absolute (Hanson, 1959; Reese, 1968). As with

many psychological debates, there may be some degree of truth on both sides. Currently, philosophers of mind (Dennett, 1991) and cognitive scientists (Kosslyn, 1980) have also wondered to what extent our mental representations and memories of the world are perfect or accurate copies of experience, and to what extent distortions (or generalizations) intrude.

The study of memory for music is potentially helpful in exploring these issues because the experimental evidence is that both relational and absolute features of music are encoded in memory. Researchers have shown that people have little trouble recognizing melodies transposed in pitch (Attneave & Olson, 1971; Dowling & Bartlett, 1981), so it is clear that memory for musical melody must encode the abstract information—that is, the relation of successive pitches, or pattern of tones if not their actual location in pitch space.

Abstract encoding has also been demonstrated for temporal features: people easily recognize songs in which the relation between rhythmic elements (the rhythmic pattern) is held constant but the overall timing or musical tempo has been changed (Monahan, 1993; Serafine, 1979). (We define tempo as the “pace” of a musical piece: measured as the amount of time it takes for a given note, or the average number of beats occurring in a given interval of time, usually beats per minute.)

With respect to this point about musical relations, Hulse and Page (1988) explain,

music emphasizes the constancy of *relations* among sounds... within *temporal* structure, music emphasizes the constancy of relations among tone durations and intertone intervals. Rhythmic structures remain perceptually equivalent over a broad range of tempos... and tempo

---

This research was supported by a National Defense Science and Engineering Graduate Fellowship to the first author, and by NSF Research Grant BNS 85-11685 to R. N. Shepard. This report was prepared in part while the first author was a Visiting Research Fellow at the Center for Computer Research in Music and Acoustics, Stanford University, Winter 1994–1995. We are grateful to the following for their generous contributions to this work: Jamshed Bharucha, Chris Chafe, Jay Dowling, Andrea Halpern, Stephen Handel, Douglas Hintzman, Jay Kadis, Carol Krumhansl, Max Mathews, Joanne Miller, Caroline Palmer, John R. Pierce, Michael Posner, Bruno Repp, Roger Shepard, Julius Smith, the researchers and staff of CCRMA, and especially Malcolm Slaney and the participants in the weekly CCRMA Hearing Seminar. Any errors remaining in this work were probably pointed out to us by these people, but we were too stubborn to change them. Correspondence may be sent to D. J. Levitin, Department of Psychology 1227, University of Oregon, Eugene, OR 97403-1227 (e-mail: levitin@cogsci.uoregon.edu). P. R. Cook is now at the Department of Computer Science, Princeton University.

changes involve ratio changes of duration and interval. (p. 431)

And, as Monahan (1993) argues, musical pitch and time are described most naturally in relative rather than absolute terms. Thus (within very broad limits), the identity and recognizability of a song are maintained through transposition of pitch and changes in tempo.

Evidence that memory for music also retains absolute pitch information alongside an abstract melody representation has been mounting for some time (Deutsch, 1991; Halpern, 1989; Levitin, 1994; Lockhead & Byrd, 1981; Terhardt & Seewann, 1983). For example, in a previous study, one of us (Levitin, 1994) asked experimental subjects (most of them nonmusicians) to sing their favorite rock and roll songs from memory, without any external reference. When the tones of the subjects' productions were compared with the tones of the recorded songs, they were found to be at or very near the songs' actual pitches. This is somewhat surprising from the standpoint of music theory (and perhaps from "object perception" theories), because that which gives a melody its identity or "objectness" is the relation of successive intervals (both rhythmic and melodic). As most music theorists would agree, "music places little emphasis on the *absolute* properties of sounds, such as their exact temporal pitch or duration" (Hulse & Page, 1988, p. 432).

The animal learning literature reveals parallel evidence for both absolute and relational memory for auditory stimuli. Starlings (Hulse & Page, 1988) and white-throated sparrows (Hurly, Ratcliffe, & Weisman, 1990) were shown to use both absolute and relative pitch information in discriminating melodies, but other evidence is mixed. Barn owls seem to remember musical patterns on the basis of absolute pitch (Konishi & Kenuk, 1975), and wolves use absolute pitch information to the exclusion of relative pitch information to differentiate between familiar and unfamiliar howls (Tooze, Harrington, & Fentress, 1990).

What reasons are there to expect that long-term memory might encode tempo information with a high degree of accuracy? Scores of conditioning experiments with animals have shown that animals can learn to estimate interval durations with great accuracy (Hulse, Fowler, & Honig, 1978).

The question of human memory for tempo has been addressed in several key studies. People report that their auditory images seem to have a specific tempo associated with them (Halpern, 1992), and in one study, subjects who imagined songs tended to imagine them at about the same tempo on occasions separated by as much as 5 days (Halpern, 1988). Collier and Collier (1994) found that trained jazz musicians tended to vary tempo less than 5% within a song or across multiple performances of the same song on different days, suggesting a stable memory for tempo in these musicians.

Whereas the studies just mentioned argue for the *stability* of internal tempos, the question of the *objective ac-*

*curacy* of these internal tempos interested us. That is, when we remember a song (or "auralize" it, to use W. D. Ward's terminology) do we do so at its original tempo? Just as some people can match pitches accurately (and we say they have "absolute pitch"), we wondered whether there are those who can match tempos accurately and have an ability of "absolute tempo." In particular, we wanted to study this ability in everyday people with little formal musical training. This question is relevant to researchers studying human rhythm perception (e.g., Desain, 1992; Jones & Boltz, 1989; Povel & Essens, 1985; Steedman, 1977), the nature and stability of internal clocks (Collier & Wright, 1995; Collyer, Broadbent, & Church, 1994; Helmuth & Ivry, 1996), mental imagery (Finke, 1985; Kosslyn, 1980), and general theories of time perception (Block, 1990; Fraisse, 1981; Michon & Jackson, 1985).

On a music-theoretic level, Narmour (1977) argues that musical listening involves using both schematic reduction and unreducible (absolute) idiostructural information. To the extent that the original listening experience is preserved in the brain, we might expect to find these two types of information represented in memory, not just in perception. One of the goals of the present study was to test this hypothesis.

## EXPERIMENT 1

Experiment 1 was designed to discover whether people encode the absolute tempo of a familiar song in memory, and if so, with what degree of precision. In Levitin (1994), subjects were asked to sing contemporary popular and rock songs from memory, and their productions were analyzed for pitch accuracy. Using these same data, we analyzed their productions for tempo accuracy. Contemporary popular and rock songs form an ideal stimulus set for our study, because they are typically encountered in only one version by a particular musical artist or group, and so the song is always heard—perhaps hundreds of times—in the same key, and at the same tempo. (In fact, we excluded songs that did not meet this criterion.)

### Method

The raw data used in this study were originally collected for a study on pitch memory (Levitin, 1994).

**Subjects.** The subjects were 46 Stanford University students who served without pay. The subjects did not know in advance that they were participating in a study involving music, and the sample included subjects with and without some musical background. All subjects filled out a general questionnaire before the experimental session. The subjects ranged in age from 16 to 35 years (mean, 19.5; mode, 18; *SD*, 3.7).

By self-report, the subjects' musical background ranged from no instruction to more than 10 years of instruction; 37 subjects had some exposure to a musical instrument, 9 had none. In response to the question "how much structured musical training in either performance or theory have you had?" 17 subjects reported none; 17 subjects reported 1–3 years; 5 subjects reported 3–5

years; 3 subjects reported 5–7 years; 3 subjects reported 7–10 years; and 1 subject reported more than 10 years.

**Materials.** Prior to data collection, a norming study was conducted to select stimuli with which this subject population would be familiar. Two hundred fifty introductory psychology students completed a questionnaire, on which they were asked to indicate songs that “they knew well and could hear playing in their heads.” None of the subjects in the norming study were subsequently used in the main experiment.

The results of this norming study were used to select the best known songs. Songs on this list that had been performed by more than one group were excluded from the stimulus set because of the possibility that these versions might have different tempos. From this questionnaire, 58 compact discs (CDs) were selected, representing over 600 songs from which the subjects could choose. Examples of songs included are “Hotel California” by The Eagles; “Get into the Groove” by Madonna; and “This & That” by Michael Penn. (A complete list of CDs constituting the stimulus set is available from the first author.)

**Procedure.** Subjects were seated in a sound attenuation booth alongside the experimenter. The 58 CDs chosen from the norming study were displayed alphabetically on a shelf in front of the subjects. The experimenter followed a written protocol asking subjects to select from the shelf and to hold in their hands a CD that contained a song they knew very well. Holding the CD and looking at it may have provided a visual cue for subsequent auditory imaging. There was no CD player in the booth, and at no time were the CDs actually played for the subjects. All subjects reported that they had not actually heard their chosen song in the previous 72 h, and many had not heard it in months.

The subjects were then asked to close their eyes and imagine that the song was actually playing. They were told that, when they were ready, we wanted them to try to reproduce the tones of the song by singing, humming, or whistling, and they could start anywhere in the song they wanted to. The subjects were not explicitly told anything about rhythm or tempo, nor were they specifically

asked to reproduce the tempo of the songs. Their productions were recorded on digital audio tape (DAT), so that the pitch and speed would be accurately preserved. The subjects were not told how much of the song to sing, but they typically sang a four-bar phrase. Following the production of this first song, the subjects were asked to choose another song and repeat the procedure; this constituted the two experimental trials. Three of the subjects discontinued their participation after Trial 1.

**Analysis.** The subjects’ productions were compared with the songs performed by the original artists on CD, in order to compare tempos. The subjects’ productions and the corresponding sections of the CD were transferred digitally to a Macintosh computer, and the sample rate was converted from its original 44.1 kHz or 48 kHz to 22.050 kHz for storage economy.

The durations of each subject’s production and the associated CD excerpt were measured with the program MacMix, and these measurements were accurate to within 0.1 msec. A typical subject production was 5 sec long. The total selection was also divided into beats yielding two equivalent measures of production time: total duration and beats per minute. For the purposes of this report, data are presented as tempos in units of beats per minute (bpm).

## Results

Figure 1 shows, as a bivariate scatterplot, the tempos produced by subjects in comparison with the actual tempos of the remembered pieces (Trials 1 and 2 are combined). The subjects came very close to their target tempos, as is indicated by the high correlation between subjects’ tempos and actual tempos ( $r = .95$ ) and the fact that most responses fall near the diagonal. Figure 2 shows the distribution of errors that the subjects made, expressed in a histogram as percent deviations from the actual tempo.<sup>1</sup>

On Trial 1, 33/46, or 72%, of the subjects performed within  $\pm 4\%$  of the actual tempo for the songs, and 41/46,

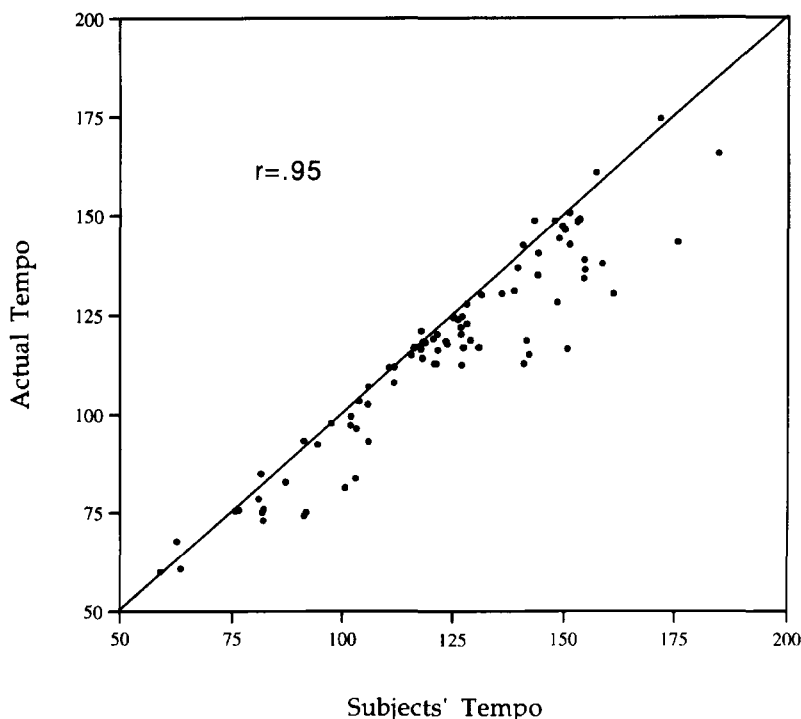


Figure 1. Subjects’ tempos versus actual tempos, both trials combined.

or 89%, of the subjects performed within  $\pm 8\%$  of the actual tempo ( $M = 4.1\%$ ;  $SD = 7.7\%$ ). On Trial 2, 12/42, or 40%, of the subjects came within  $\pm 4\%$ , and 25/42, or 60%, came within  $\pm 8\%$  of the actual tempo ( $M = 7.7\%$ ;  $SD = 7.9\%$ ). As Figure 2 shows, for the two trials combined, 72% of the responses fell within 8%.

To put these results in context, one might ask, what is the JND for tempo? Drake and Botte (1993, Experiment 3) found the JND for tempo discrimination to be 6.2%–8.8% using a two-alternative forced choice listening (“which is faster?”) test; Friberg and Sundberg (1993) found JND for tempo to be 4.5% using the psychophysical method of adjustment; Hibi (1983) found the JND to be  $\sim 6\%$  for displacement of a single time marker in a sequence, and for lengthening/shortening of a single time marker. In tapping tasks, where subjects had to either tap along with a pulse at a certain tempo (a “synchronization task”) or continue tapping to a tempo set up by the experimenter (“continuation task”), JNDs of 3%–4% have been reported for synchronization (Collyer et al., 1994; Povel, 1981), and 7%–11% for continuation (Allen, 1975). All of these JNDs apply for tempos in the range our subjects sang. It appears, then, that a large percentage of the subjects in our study performed within one or two JNDs for tempo, based solely on their *memory* for the musical pieces.

A somewhat more ecologically valid confirmation of these JND figures comes from Perron (1994). In contemporary popular music, many recordings are made with drum machines or computer sequencers instead of live drummers. The anecdotal opinion of musicians and record producers has been that these machines are much more able to hold a steady tempo than are human players. Perron measured the tempo deviations in a number of these devices and found the mean tempo deviation to be 3.5% (with a standard deviation of 4.5%). Because the deviations in these machines seem to go largely un-

noticed by most people (including professional drummers), it seems fair to assume that this 3.5% is less than the JND for tempo variation. An interesting implication of Perron’s finding is that our subjects may well have been trying to reproduce tempos for songs that contain variations of this magnitude.

One might ask whether the subjects in our study performed consistently across the two trials. To measure this, trials on which the subjects came within  $\pm 6\%$  were considered “hits” and all others were considered “misses,” in accordance with the more conservative of Drake and Botte’s (1993) JND estimates. Twenty-five subjects (or 60%) were found to be consistent in their performance across trials. Yule’s  $Q$  was computed as a measure of strength of association for this  $2 \times 2$  table, and was significant ( $Q = .50$ ;  $p < .04$ ).

Next, we wondered whether the subjects who had accurate tempo memory also had accurate pitch memory (as measured in Levitin, 1994). Combining Trials 1 and 2, and using 6% as a “hit” criterion for tempo and  $\pm 1$  semitone (s.t.) as a criterion for pitch, we did not find a significant association (Yule’s  $Q = .31$ ; n.s.).<sup>2</sup>

One of the implicit assumptions in these analyses is that the tempos of the songs that our subjects sang were widely distributed. If all the songs fell into a narrow tempo band, one might argue that the subjects had memory for only a particular tempo (or narrow set of tempos). As Figure 1 shows, however, the range of tempos produced by subjects was very large, running from approximately 60 bpm to over 160 bpm.

Similarly, one might wonder whether the good performance of subjects across trials was merely due to all of the subjects’ singing songs at the same tempo in both cases; individual subjects may have an idiosyncratic “preferred” tempo (or “internal tempo”) that they knew well and relied on for this task (Braun, 1927). We found that the correlation between tempos sung on Trial 1 and

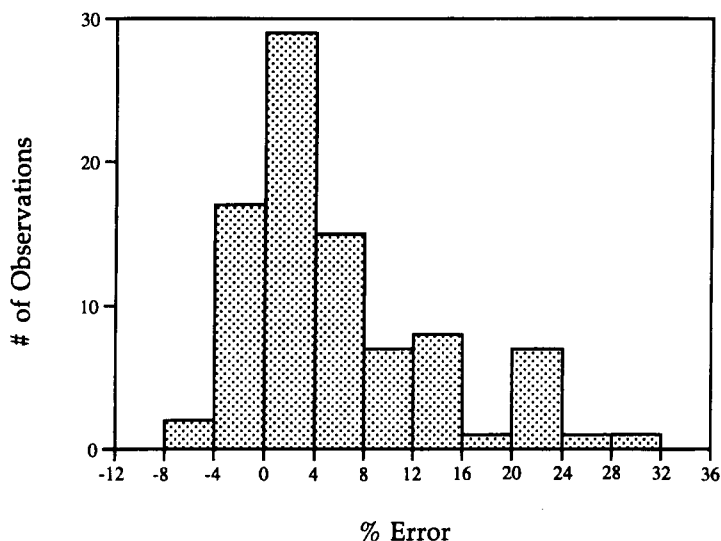


Figure 2. Percent deviation from actual tempo, both trials combined.

Trial 2 was very low ( $r = .07$ ), as was the correlation between “targets” on Trial 1 and Trial 2 (the tempos subjects were trying to reproduce;  $r = .04$ ).

Variations of  $\pm 5\%$  during the subjects’ learning of the material could have occurred if subjects had heard the songs repeatedly on a cassette player or record player that did not keep accurate speed; CD players are not subject to speed fluctuations. In one item on the questionnaire, the subjects were asked about whether they had heard their chosen song on CD, radio, cassette, or record player, and no correlation was found between the source of the learning and their performance (all the commercial radio stations in the area of the study were broadcasting CDs exclusively during the study period, so “radio” responses were considered to be “CD learning”). Similarly no correlation was found between other factors such as sex, age, handedness, or musical training.

### Discussion

The finding that 89% of the Trial 1 subjects and 60% of the Trial 2 subjects made errors of only  $\pm 8\%$  is evidence that long-term memory for tempo is very accurate and is near the discrimination threshold (as measured by JNDs) for variability in tempo. These results may even underestimate the strength of tempo memory, because our subjects were only instructed to reproduce *pitches* accurately; to the extent that they also reproduced tempo, they did this on their own, without being requested by the experimenter to do so.

The distribution of errors is also instructive. As Figures 1 and 2 show, there is a tight clustering near the actual tempo, and more subjects sang too fast than too slow. Boltz (1994) reviews evidence that various forms of induced stress increase the internal tempo of individuals. This would create internal durations that are shorter than the standard and would cause the subjects to sing fast. If we can assume that the experimental situation was somewhat stressful (many subjects seemed to be embarrassed or nervous), this could account for the asymmetric error distribution favoring faster reproductions. An additional explanation for this asymmetry comes from experimental findings that people are more likely to perform faster rather than slower (Kuhn, 1977) and are better able to detect tempo decreases rather than increases (Kuhn, 1974).

In spite of the certain awkwardness of being asked to sing for a psychological experiment and the concomitant desire to be done with the task as quickly as possible, most subjects reproduced the tempo of their selected songs with remarkable accuracy. In listening to the subjects’ productions and to the corresponding artists’ renditions, we were struck by how close the subjects came not just in tempo, but in pitch, phrasing, and stylistic nuances while singing from memory. In many cases, it seemed that the subjects could not have performed better if they were actually singing along with the CD—but of course, they were merely singing along with a representation of the CD in their heads.

One could argue that our findings are the result of an artifact rather than actual “memory for tempo.” While recalling these musical pieces, the subjects sing or imagine lyrics, and perhaps the lyrics provide a constraint for the tempo. That is to say, the tempo of a piece might be constrained by the number of syllables that have to be fit into a particular melody. To counter this argument, one would need to find a piece of music with lyrics that had no well-defined tempo standard and ask subjects to sing it. If the range of tempos produced for such a song is much wider than the range produced by our subjects, we could argue that articulatory constraints do not account for memory for tempo.

As a first look at this issue, we examined Halpern’s (1988) data. In her Experiment 1, Halpern asked subjects to imagine popular songs and set a metronome to match the tempo that they heard in their heads. Most of the songs had the property that no single reference (or canonical) version existed—for example, “Happy Birthday,” “London Bridge is Falling Down,” and “Twinkle, Twinkle, Little Star.” In general, peoples’ exposure to these songs occurs through informal live singing (such as in elementary school), and the variety of recorded versions of these songs virtually ensures that there is no uniformly “correct” tempo. Halpern provided us with the (previously unpublished) standard deviations across subjects for these three songs, and they were (respectively) 16%, 19%, and 22% (see Table 1).

In a replication (Experiment 2), Halpern (1988) asked a second group of subjects to perform the same task. The mean tempo (across subjects) for each song varied by a large amount from one experiment to the other: for the three songs (respectively), the differences in mean tempos were 19%, 12%, and 14%. Halpern also asked her Experiment 2 subjects to adjust a metronome incrementally to the point that represented the fastest and slowest that they could imagine the song to be. As Table 1 indicates, the tempos selected by the subjects varied more than 250%. All of Halpern’s tempo variations are larger than those we found in our subjects, supporting our claim that the tempo was not tightly constrained by the lyrics.

As a control condition, we designed Experiment 2 in order to replicate Halpern’s (1988) earlier findings about tempo variability in a production context. We recruited 8 subjects and asked them to sing three familiar folk songs (mentioning nothing to them about tempo), and afterward to sing them as fast and slow as possible. A large standard deviation in this task would replicate Halpern’s finding and indicate that lyrics do not significantly constrain tempo.

## EXPERIMENT 2

### Method

**Subjects.** The subjects were 4 University of Oregon students, and 4 members of the community, recruited without regard to musical training; 6 subjects had no previous musical instruction, 2

**Table 1**  
**Song Tempos From Halpern (1988)**

Title	Experiment 1		Experiment 2	% Difference	Experiment 2		Fast to Slow Difference
	<i>M</i>	<i>SD</i>			Fast	Slow	
"Happy Birthday"	101	16%	120	19%	163	65	251%
"London Bridge"	101	19%	113	12%	170	65	262%
"Twinkle Twinkle"	99	22%	113	14%	168	63	267%

Note—Tempos are given in beats per minute. % Difference is for Experiment 1 minus Experiment 2. Adapted, with permission, from "Perceived and Imagined Tempos of Familiar Songs," by A. R. Halpern, 1988, *Music Perception*, 6, p. 198. Copyright 1988 by the Regents of the University of California. This table also includes new, previously unpublished information supplied by the author, as well as our own analyses of the data.

subjects had less than 2 years of musical instruction. All the subjects served without pay.

**Materials.** To investigate tempo variability in singing, the subjects were asked to sing "Happy Birthday," "We Wish You A Merry Christmas," and "Row, Row, Row Your Boat."

**Procedure.** The subjects were asked to sing one of the three songs (song order was randomized). When they finished, they were next asked to sing it "as slow as you possibly can" and then "as fast as you possibly can." This was repeated for the other two songs. The subjects were recorded either directly to the hard disk of a NeXT computer using the program SoundEditor, or to a Sony DATMAN digital audio tape recorder which was then transferred digitally to NeXT computer sound files.

## Results

Table 2 shows the distribution of tempos for the three songs sung at their normal speeds. The standard deviations are all well above the ~8% standard deviation that we found for our Experiment 1 subjects. An *F* test between the variance in Experiment 1 and the lowest of the Experiment 2 variances (for "Row, Row, Row Your Boat") confirmed that the variances were significantly different [ $F(1,7) = 24.83, p < .01$ ], using Levene's test for homogeneity of variances, and Satterthwaite's correction for unequal *n* (Snedecor & Cochran, 1989).

Figure 3 shows the variability of tempos expressed as percent deviation from each song's mean tempo. When compared with Figure 2, the greater variability in tempos is easy to see. This replicates Halpern's (1988) finding that production variability on these types of songs is large. Furthermore, the fast and slow performances of each of the three songs showed that there is indeed a large range over which people can produce familiar songs with lyrics. The song "Happy Birthday" exhibited maxima and minima of 421 and 48 bpm (with the mean across subjects being 284 and 76 bpm). "We Wish You A Merry Christmas" exhibited maxima and minima of 129 and 22 (with means of 102 and 36). "Row, Row, Row

Your Boat" exhibited maxima and minima of 280 and 41 (with means of 226 and 72). This large deviation in "normal" speeds, coupled with the large range of possible speeds, confirms that a given song can be sung across a very broad range of tempos, and that lyric or articulatory constraints were probably not playing a role in our Experiment 1 subjects' accurate memory for tempo.

## Discussion

In Experiment 2, we replicated Halpern's (1988) finding that the variability in tempos for popular songs that lack a tempo standard is in the 10%–20% range, well exceeding the variability of our Experiment 1 subjects. The accurate performance in Experiment 1 does not seem to have been due to constraints imposed by lyrics.

## GENERAL DISCUSSION

Songs contain both pitch and tempo information during their performance. What can we say about the mental representation of songs in the brain? Drake and Botte (1993) argued for the existence of a single brain mechanism that judges the tempo of sequences (not merely the durations of intervallic events). Judgment of tempos might be controlled by a central timing mechanism located in the cerebellum (Helmuth & Ivry, 1996), the operation of which is based on oscillatory processes (Ivry & Hazeltine, 1995). Such an "internal clock" may not keep perfect time, but be subject to  $1/f$  noise (Gilden, Thornton, & Mallon, 1995). But pitch perception seems to occur in brain systems separate from time perception, beginning with frequency-selective cells in the cochlea, all the way through to the auditory cortex (Moore & Glasberg, 1986). So it would seem that the perception of pitch and tempo is handled by different systems.

But memory for songs may somehow combine or link pitch and tempo representations. Our intuition—despite the finding that there was not a statistically significant correlation between tempo memory and pitch memory—is that the entire spectral-temporal profile of a song is encoded in memory in some fashion and that repeated listenings strengthen the trace. Pitt (1995) provided evidence for a central representation of instrumental timbre, which supports the notion that memory preserves a complex, spectral-temporal image.

**Table 2**  
**Means and Standard Deviations for Experiment 2**

Song Title	<i>M</i> (bpm)	<i>SD</i> (bpm)	<i>SD</i> (as percent)
"Happy Birthday"	137	29.1	21%
"We Wish You A Merry Christmas"	63	11.7	18%
"Row, Row, Row Your Boat"	134	16.2	12%

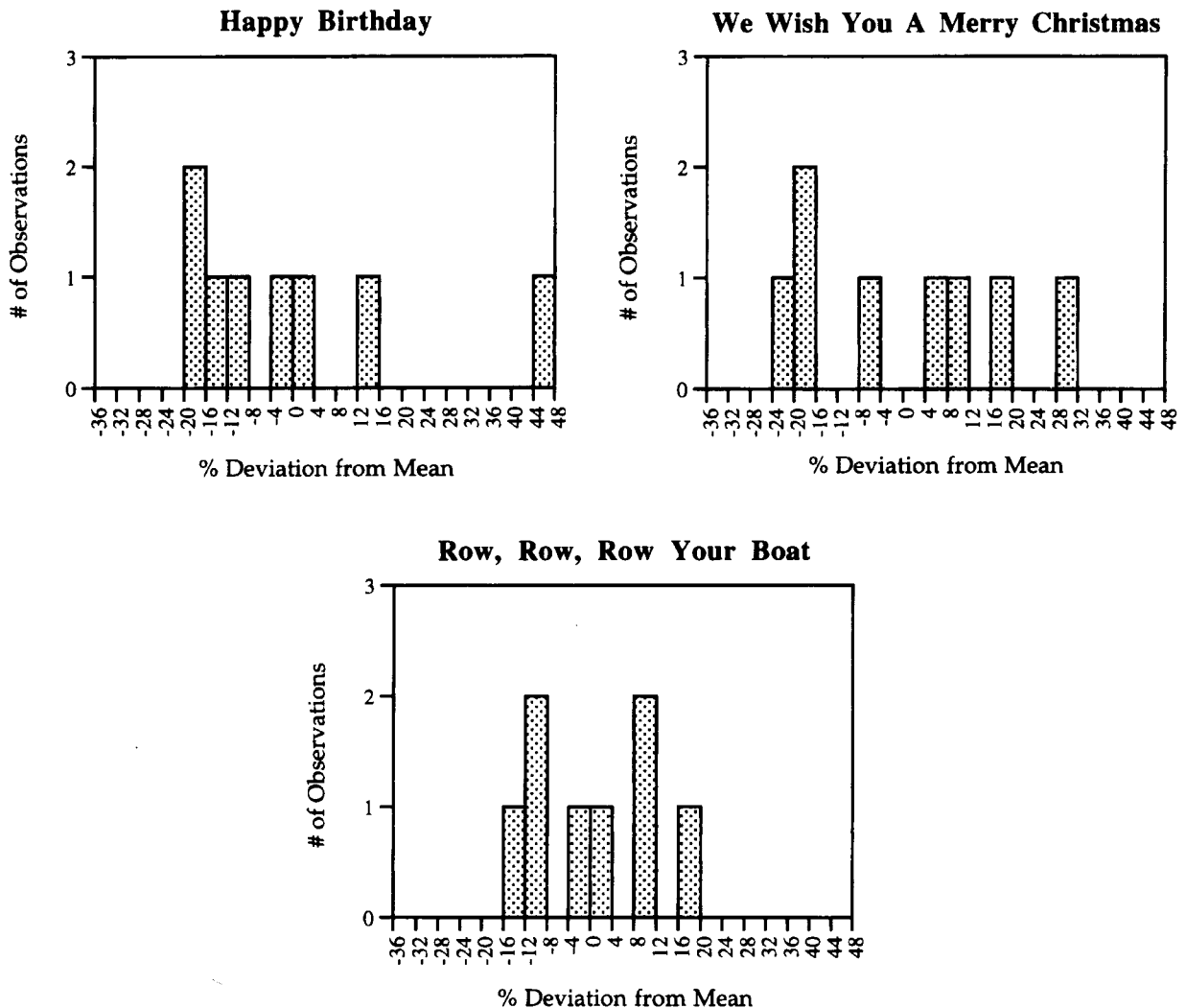


Figure 3. Distribution of tempos in Experiment 2. For "Happy Birthday,"  $SD = 21\%$ ; for "We Wish You A Merry Christmas,"  $SD = 18\%$ ; for "Row, Row, Row Your Boat,"  $SD = 12\%$ .

In any event, it seems increasingly clear that human memory encodes both the absolute and the relative information contained in musical pieces, and that people are able to access whichever is required by the given task. This supports previous theoretical predictions that memory does encode absolute features of the original stimulus, along with abstract relations (Bower, 1967; Hintzman, 1986). This would also account for peoples' ability to easily recognize songs in transposition and for our findings of their being able to reproduce a particular absolute feature. Premack (1978) offers an account of the relation between abstract and absolute memory, suggesting that abstraction is induced only as a response to an overburdened memory; that is, only absolute cues are memorized until memory becomes taxed, and then the organism forms an abstract rule.

Some colleagues have wondered whether our results are merely the effect of "overlearning" of the stimuli, and they have suggested that our findings are nothing

more than a measure of how well the stimuli were learned. But we agree with Palmer (S. E. Palmer, personal communication, October 1994), who argues that increased learning is just a matter of increasing the signal-to-noise ratio in memory retrieval. To toss aside the present findings as "merely overlearning" is to miss our point; to paraphrase Halpern (1992), we are interested in the nature of what is encoded in memory when memory is working well. Overlearned or well-learned stimuli provide the cleanest measure of this.

It is well established, at least anecdotally, that expert musicians are capable of producing tempos from memory with great accuracy. In the present study, we found that even nonmusicians have an accurate representation of tempo that they are able to reproduce. Considering this, together with the results of Levitin (1994) and other studies, we wonder (perhaps somewhat facetiously) whether what people encode in memory is the first bar of what would be written music, including the key, time sig-

nature, and metronome marking! This would be parsimonious, allowing the brain to store only the temporal and melodic relations between tones and imposing temporal and pitch anchors only when needed. And it would explain the ease with which people are able to identify transpositions of pitch and changes in tempo. But our subjective impression, based on introspection, is that *long-term* memory for music functions more like what Attneave and Olson (1971) described as *short-term* music memory:

The circulating short-term memory of a tonal sequence just heard, which is experienced as an auditory image extended in real time (as a melody that "runs through one's head"), typically preserves the key or specific pitch values of the original. (p. 164)

Attneave and Olson (1971) believed that, in contrast, the long-term memory trace is encoded only on the basis of relations, or intervals. They note that there is nothing about such a coding system that precludes the additional storage of pitch or tempo information, "but it is evident that normal individuals do not, in fact, preserve this kind of information with any high degree of precision" (pp. 164-165). We believe that the present study provides evidence against this point, and that ordinary individuals *do* possess representations that are more accurate than has previously been believed. Furthermore, the present study provides evidence that the two types of operations present in musical listening—schematic reduction and unreducible idiostructural information (Narmour, 1977)—appear to also be present in long-term memory.

## REFERENCES

- ALLEN, G. D. (1975). Speech rhythm: Its relation to performance universals and articulatory timing. *Journal of Phonetics*, **3**, 75-86.
- ATTNEAVE, F., & OLSON, R. K. (1971). Pitch as a medium: A new approach to psychophysical scaling. *American Journal of Psychology*, **84**, 147-166.
- BLOCK, R. A. (ED.) (1990). *Cognitive models of psychological time*. Hillsdale, NJ: Erlbaum.
- BOLTZ, M. G. (1994). Changes in internal tempo and effects on the learning and remembering of event durations. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 1154-1171.
- BOWER, G. H. (1967). A multicomponent theory of the memory trace. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 1, pp. 229-325). New York: Academic Press.
- BRAUN, F. (1927). Untersuchungen über das persönliche tempo [Investigations into personal tempo]. *Archiv der gesamten Psychologie*, **60**, 317-360.
- BROOKS, L. R. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 169-211). Hillsdale, NJ: Erlbaum.
- COLLIER, G. L., & COLLIER, J. L. (1994). An exploration of the use of tempo in jazz. *Music Perception*, **11**, 219-242.
- COLLIER, G. L., & WRIGHT, C. E. (1995). Temporal rescaling of simple and complex ratios in rhythmic tapping. *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 602-627.
- COLLYER, C. E., BROADBENT, H. A., & CHURCH, R. M. (1994). Preferred rates of repetitive tapping and catagorical time production. *Perception & Psychophysics*, **55**, 443-453.
- CROWDER, R. G., SERAFINE, M. L., & REPP, B. (1990). Physical interaction and association by contiguity in memory for the words and melodies of songs. *Memory & Cognition*, **18**, 469-476.
- DENNETT, D. C. (1991). *Consciousness explained*. Boston: Little, Brown.
- DESAIN, P. (1992). A (de)composable theory of rhythm perception. *Music Perception*, **9**, 439-454.
- DEUTSCH, D. (1991). The tritone paradox: An influence of language on music perception. *Music Perception*, **8**, 335-347.
- DOWLING, W. J., & BARTLETT, J. C. (1981). The importance of interval information in long-term memory for melodies. *Psychomusicology*, **1**, 30-49.
- DRAKE, C., & BOTTE, M.-C. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. *Perception & Psychophysics*, **54**, 277-286.
- FINKE, R. A. (1985). Theories relating mental imagery to perception. *Psychological Bulletin*, **98**, 236-259.
- FRAISSE, P. (1981). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 149-180). New York: Academic Press.
- FRIBERG, A., & SUNDBERG, J. (1993). Perception of just noticeable time displacement of a tone presented in a metrical sequence at different tempos. In Kungliga Tekniska Högskolan, Speech Transmission Laboratory [Eds.], *Quarterly progress and status report* (pp. 49-55). Stockholm: Royal Institute of Technology, Department of Speech Communication, Speech Transmission Laboratory.
- GILDEN, D. L., THORNTON, T., & MALLON, M. W. (1995, March 24). 1/f noise in human cognition. *Science*, **272**, 1837-1839.
- HALPERN, A. R. (1988). Perceived and imagined tempos of familiar songs. *Music Perception*, **6**, 193-202.
- HALPERN, A. R. (1989). Memory for the absolute pitch of familiar songs. *Memory & Cognition*, **17**, 572-581.
- HALPERN, A. R. (1992). Musical aspects of auditory imagery. In D. Reisberg (Ed.), *Auditory imagery* (pp. 1-28). Hillsdale, NJ: Erlbaum.
- HANSON, H. M. (1959). Effects of discrimination training on stimulus generalization. *Journal of Experimental Psychology*, **58**, 331-334.
- HELMUTH, L. L., & IVRY, R. B. (1996). When two hands are better than one: Reduced timing variability during bimanual movements. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 278-293.
- HIBI, S. (1983). Rhythm perception in repetitive sound sequence. *Journal of the Acoustical Society of Japan*, **4**, 83-95.
- HINTZMAN, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, **93**, 411-428.
- HULSE, S. H., FOWLER, H., & HONIG, W. K. (EDS.) (1978). *Cognitive processes in animal behavior*. Hillsdale, NJ: Erlbaum.
- HULSE, S. H., & PAGE, S. C. (1988). Toward a comparative psychology of music perception. *Music Perception*, **5**, 427-452.
- HURLY, T. A., RATCLIFFE, L., & WEISMAN, R. (1990). Relative pitch recognition in white-throated sparrows, *Zonotrichia albicollis*. *Animal Behavior*, **40**, 176-181.
- IVRY, R. B., & HAZELTINE, R. E. (1995). The perception and production of temporal intervals across a range of durations: Evidence for a common timing mechanism. *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 3-18.
- JACOBY, L. L. (1983). Remembering the data: Analyzing interaction processes in reading. *Journal of Verbal Learning & Verbal Behavior*, **22**, 485-508.
- JONES, M. R., & BOLTZ, M. (1989). Dynamic attending and responses to time. *Psychological Review*, **96**, 459-491.
- KONISHI, M., & KENUK, A. S. (1975). Discrimination of noise spectra by memory in the barn owl. *Journal of Comparative Physiology*, **97**, 55-58.
- KOSSLYN, S. M. (1980). *Image and mind*. Cambridge, MA: Harvard.
- KUHN, T. L. (1974). Discrimination of modulated beat tempo by professional musicians. *Journal of Research in Music Education*, **22**, 270-277.
- KUHN, T. L. (1977). Effects of dynamics, halves of exercise, and trial sequences on tempo accuracy. *Journal of Research in Music Education*, **25**, 222-227.
- LEVITIN, D. J. (1994). Absolute memory for musical pitch: Evidence from the production of learned melodies. *Perception & Psychophysics*, **56**, 414-423.
- LOCKHEAD, G. R., & BYRD, R. (1981). Practically perfect pitch. *Journal of the Acoustical Society of America*, **70**, 387-389.



- MCCLELLAND, J. L., & RUMELHART, D. E. (1986). A distributed model of human learning and memory. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel distributed processing: Vol. 2. Psychological and biological models* (pp. 170-215). Cambridge, MA: MIT Press.
- MEDIN, D. L., & SCHAFFER, M. M. (1978). Context theory of classification learning. *Psychological Review*, **85**, 207-238.
- MICHON, J. A., & JACKSON, J. L. (Eds.) (1985). *Time, mind, and behavior*. New York: Springer-Verlag.
- MONAHAN, C. B. (1993). Parallels between pitch and time and how they go together. In T. J. Tighe & W. J. Dowling (Eds.), *Psychology and music: The understanding of melody and rhythm* (pp. 121-154). Hillsdale, NJ: Erlbaum.
- MOORE, B. C. J., & GLASBERG, B. R. (1986). The relationship between frequency selectivity and frequency discrimination for subjects with unilateral and bilateral cochlear impairment. In B. C. J. Moore & R. D. Patterson (Eds.), *Auditory frequency selectivity* (pp. 407-414). New York: Plenum.
- NARMOUR, E. (1977). *Beyond Schenkerism: The need for alternatives in music analysis*. Chicago: University of Chicago Press.
- PERRON, M. (1994, November). *Checking tempo stability of MIDI sequencers*. Paper presented at the 97th Convention of the Audio Engineering Society, San Francisco.
- PITT, M. A. (1995). Evidence for a central representation of instrument timbre. *Perception & Psychophysics*, **57**, 43-55.
- POSNER, M. I., & KEELE, S. W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology*, **83**, 304-308.
- POVEL, D. J. (1981). Interval representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 3-18.
- POVEL, D. J., & ESSENS, P. (1985). Perception of temporal patterns. *Music Perception*, **2**, 411-440.
- PREMACK, D. (1978). On the abstractness of human concepts: Why it would be difficult to talk to a pigeon. In S. H. Hulse, H. Fowler, & W. K. Honig (Eds.), *Cognitive processes in animal behavior* (pp. 423-451). Hillsdale, NJ: Erlbaum.
- REESE, H. W. (1968). *The perception of stimulus relations*. New York: Academic Press.
- ROSCH, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, **104**, 192-223.
- SERAFINE, M. L. (1979). A measure of meter conservation in music, based on Piaget's theory. *Genetic Psychology Monographs*, **99**, 185-229.
- SNEDECOR, G. W., & COCHRAN, W. G. (1989). *Statistical methods* (8th ed.). Ames, IA: Iowa State University Press.
- STEEDMAN, M. J. (1977). The perception of musical rhythm and metre. *Perception*, **6**, 555-570.
- TERHARDT, E., & SEEWAN, M. (1983). Aural key identification and its relationship to absolute pitch. *Music Perception*, **1**, 63-83.
- TOOZE, Z. J., HARRINGTON, F. H., & FENTRESS, J. C. (1990). Individually distinct vocalizations in timber wolves, *Canis lupus*. *Animal Behavior*, **40**, 723-730.

## NOTE

1. Only 88 data points are shown in Figures 1 and 2. We started with 46 subjects in Trial 1; 3 discontinued participation after Trial 1, and we eliminated the data of one Trial 2 subject. This Trial 2 subject asked to sing a Mozart piano sonata on Trial 2, complaining that she didn't know any rock songs other than her Trial 1 choice ("Hey Jude"). The experimenter let her sing the piano sonata. Later, during the analysis phase of the study, we realized that multiple recorded versions of such a piece exist, and virtually all are in the same key, so this subject's production was not excluded from the original pitch analysis reported in Levitin (1994). However, the range of tempos over which such pieces are performed typically varies, so this subject was excluded in the tempo analysis, on the ground that a single reference standard did not exist.

2. The low correlation between pitch memory and tempo memory could indicate that subjects have either *independent* storage or *integrated* storage of these two attributes, in the sense of the terms proposed by Crowder, Serafine, and Repp (1990). In independent storage, memory for one element is uninfluenced by the other; in integrated storage, the elements are related in memory in such a way that one component is better recognized in the presence of the other than in its absence. The present data do not provide direct evidence for choosing between these two possibilities, but it is our intuition that tempo and pitch are best characterized by an integrated storage account. Some subjects are able to recall both attributes with great accuracy (one attribute might indeed aid the accurate recall of the other) while other subjects showed no such benefit.

(Manuscript received July 11, 1995;  
revision accepted for publication December 14, 1995.)